

## Claims

What is claimed is:

1. A method for transforming gene expression signals, the method  
5 comprising the steps of:  
determining a plurality of gene expression signals for a gene; and  
deriving a transformation that renders uniform, within a selected interval, a  
distribution of transformed gene expression signals for the gene.
- 10 2. The method of claim 1, further comprising the step of applying the  
transformation to an additional gene expression signal.
3. The method of claim 1, wherein the step of deriving comprises the steps  
of:  
15 determining a function that approximates the distribution; and  
using the function to create the transformation, wherein the transformation  
renders uniform a probability distribution of the transformed gene expression signals, and  
wherein each gene expression signal is mapped by the transformation into a transformed  
gene expression signal.

20

4. A method for the characterization of phenotypes by gene expression patterns and classification of samples based thereon, the method comprising the steps of:

a) forming, from a plurality of first genes and first samples of a first phenotype, a first matrix of gene expression signals for the first phenotype, each gene expression signal corresponding to one of the first genes and one of the first samples;

b) forming, from a plurality of selected genes and selected samples of a selected phenotype, a second matrix of gene expression signals for the selected phenotype, each gene expression signal corresponding to one of the selected genes and one of the selected samples, wherein each selected gene corresponds to one of the first genes;

c) deriving a plurality of transformations, wherein each transformation renders uniform, within a selected interval, a distribution of transformed gene expression signals for one of the first genes, whereby each transformation corresponds to a selected gene;

d) applying, for each of the selected genes, one of the transformations to the gene expression signals of a corresponding one of the selected genes of the second matrix; and

e) defining at least one submatrix of the second matrix, wherein each transformed gene expression signal for each gene of the submatrix falls within a selected range.

5. The method of claim 4, wherein a number of first genes is the same as a number of selected genes.

6. The method of claim 4, wherein a number of selected genes is less than a number of first genes.

7. The method of claim 4, wherein each of the at least one submatrices is a largest distinct submatrix that satisfies step (e).

8. The method of claim 7, wherein the at least one submatrix comprises a plurality of submatrices.

9. The method of claim 4, wherein step (c) comprises the steps of, for each of the transformations:

i) determining a function that approximates the distribution of expression levels for a corresponding one of the first genes; and

ii) using the function to create the transformation corresponding to the one first gene, wherein the transformation renders uniform a probability distribution of the transformed gene expression signals, and wherein each gene expression signal is mapped by the transformation into a transformed gene expression signal.

10. The method of claim 9, wherein step (d) comprises the steps of:

i) selecting a gene expression signal in the second matrix;

ii) selecting one transformation that corresponds to the gene expression signal;

iii) determining a transformed gene expression signal by using the one transformation and the gene expression signal; and

iv) performing the preceding three steps, selecting another gene expression signal each time, until all of the gene expression signals have been selected and their respective transformed values determined.

11. The method of claim 4, wherein each of steps (a) and (b) comprises a step of depositing biological material from an organism upon a gene expression array.

12. The method of claim 11, wherein the biological material comprises a marker substance operable to exhibit a perceptible gene expression signal for indicating a relationship between a gene element and the biological material.
- 5 13. The method of claim 12, wherein the marker substance is operable to exhibit fluorescence as the perceptible gene expression signal.
14. The method of claim 4, further comprising the steps of:
- 10 f) determining a gene expression signal from biological material of an organism for each gene of each of the at least one submatrices; and
- g) determining a deviation of each of the gene expression signals of the organism from the selected range of each selected gene in the at least one submatrix.
15. The method of claim 14, further comprising the step of:
- 15 h) determining if the deviation of each of the gene expression signals from the biological material of the organism meets predetermined criteria, whereby the organism exhibits the selected phenotype if the deviation of each of the gene expression signals meets the predetermined criteria.
- 20 16. The method of claim 15:
- wherein the at least one submatrix comprises a plurality of submatrices;
- wherein step (f) comprises the step of forming an organism vector from the gene expression signals from the biological material of the organism;
- wherein step (g) comprises the steps of:
- 25 i) selecting at least one of the plurality of submatrices;
- ii) converting each of the at least one selected submatrices to a vector;
- iii) using the vectors for each of the selected submatrices in a

discriminant function to compare the vectors from the submatrices to the organism vector; and

wherein step (h) comprises the step of determining if the discriminant function returns a value greater than a threshold.

5

17. A system comprising:

a memory that stores computer-readable code; and

a processor operatively coupled to the memory, the processor configured to implement the computer-readable code, the computer-readable code configured to:

10 determine a plurality of gene expression signals for a gene; and

derive a transformation that renders uniform, within a selected interval, a distribution of transformed gene expression signals for the gene.

18. The system of claim 17, wherein the computer-readable code is further  
15 configured to apply the transformation to an additional gene expression signal.

19. The system of claim 17, wherein the computer-readable code is further configured, during the step of deriving, to perform the steps of:

determine a function that approximates the distribution; and

20 use the function to create the transformation, wherein the transformation renders uniform a probability distribution of the transformed gene expression signals, and wherein each gene expression signal is mapped by the transformation into a transformed gene expression signal.

25 20. A system comprising:

a memory that stores computer-readable code; and

a processor operatively coupled to the memory, the processor configured

to implement the computer-readable code, the computer-readable code configured to:

form, from a plurality of first genes and first samples of a first phenotype, a first matrix of gene expression signals for the first phenotype, each gene expression signal corresponding to one of the first genes and one of the first samples;

5 form, from a plurality of selected genes and selected samples of a selected phenotype, a second matrix of gene expression signals for the selected phenotype, each gene expression signal corresponding to one of the selected genes and one of the selected samples, wherein each selected gene corresponds to one of the first genes;

10 derive a plurality of transformations, wherein each transformation renders uniform, within a selected interval, a distribution of transformed gene expression signals for one of the first genes, whereby each transformation corresponds to a selected gene;

apply, for each of the selected genes, one of the transformations to the gene expression signals of a corresponding one of the selected genes of the second matrix; and

15 define at least one submatrix of the second matrix, wherein each transformed gene expression signal for each gene of the submatrix falls within a selected range.

21. The system of claim 20, wherein the computer-readable code is further  
20 configured to:

determine a gene expression signal from biological material of an organism for each gene of each of the at least one submatrices; and

25 determine a deviation of each of the gene expression signals from the biological material of the organism from the selected range of each selected gene in the at least one submatrix.

22. The system of claim 20, wherein the computer-readable code is further

configured to:

determine if the deviation of each of the gene expression signals from the biological material of the organism meets predetermined criteria, whereby the organism exhibits the selected phenotype if the deviation of each of the gene expression signals meets the predetermined criteria.

23. An article of manufacture comprising:

a computer readable medium having computer readable code means embodied thereon, the computer readable program code means comprising:

a step to determine a plurality of gene expression signals for a gene; and  
a step to derive a transformation that renders uniform, within a selected interval, a distribution of transformed gene expression signals for the gene.

24. The article of manufacture of claim 23, wherein the computer-readable code means further comprises a step to apply the transformation to an additional gene expression signal.

25. The article of manufacture of claim 23, wherein the computer-readable code means is further configured, during the step of deriving, to perform:

a step to determine a function that approximates the distribution; and  
a step to use the function to create the transformation, wherein the transformation renders uniform a probability distribution of the transformed gene expression signals, and wherein each gene expression signal is mapped by the transformation into a transformed gene expression signal.

26. An article of manufacture comprising:

a computer readable medium having computer readable code means embodied thereon, the computer readable program code means comprising:

5 a step to form, from a plurality of first genes and first samples of a first phenotype, a first matrix of gene expression signals for the first phenotype, each gene expression signal corresponding to one of the first genes and one of the first samples;

10 a step to form, from a plurality of selected genes and selected samples of a selected phenotype, a second matrix of gene expression signals for the selected phenotype, each gene expression signal corresponding to one of the selected genes and one of the selected samples, wherein each selected gene corresponds to one of the first genes;

15 a step to derive a plurality of transformations, wherein each transformation renders uniform, within a selected interval, a distribution of transformed gene expression signals for one of the first genes, whereby each transformation corresponds to a selected gene;

a step to apply, for each of the selected genes, one of the transformations to the gene expression signals of a corresponding one of the selected genes of the second matrix; and

20 a step to define at least one submatrix of the second matrix, wherein each transformed gene expression signal for each gene of the submatrix falls within a selected range.

27. The article of manufacture of claim 26, wherein the computer-readable code means further comprises:

25 a step to determine a gene expression signal from biological material of an organism for each gene of each of the at least one submatrices; and

a step to determine a deviation of each of the gene expression signals from



the biological material of the organism from the selected range of each selected gene in the at least one submatrix.

28. The article of manufacture of claim 26, wherein the computer-readable  
5 code means further comprises:

a step to determine if the deviation of each of the gene expression signals from the biological material of the organism meets predetermined criteria, whereby the organism exhibits the selected phenotype if the deviation of each of the gene expression signals meets the predetermined criteria.

10

For office use only